DOCUMENT RESUME

ED 304 116 IR 013 675

AUTHOR Hunka, S.
TITLE Fifteen Years of Teaching Elementary Applied Statistics Using CAI.
INSTITUTION Alberta Univ., Edmonton. Div. of Educational Research Services.
REPORT NO RIR-88-6
PUB DATE 7 Nov 88
NOTE 21p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Computer Assisted Instruction; *Course Content; *Courseware; Foreign Countries; Higher Education; *Instructional Design; Mathematics Instruction; *Microcomputers; Statistics
IDENTIFIERS *Mainframe Computers; *University of Manitoba (Canada)

ABSTRACT

A computer-assisted instructional (CAI) course in applied statistics has been taught for 15 years in the Faculty of Education at the University of Manitoba. The CAI courseware was originally created to be the primary mode of instruction for the course, and it is very extensive in terms of content and style of presentation. The course includes 14 modules of instruction and 10 examinations, and it takes the average student about 80-90 hours of online instruction to complete. Originally programmed in IBM's Coursewriter II authoring language for use on an IBM 1500 system, the course continues to exist in this language, with some enhancement provided through the development of an in-house interpreter. Under the present CAI system, the course requires about 2.3 megabytes of memory, not counting the memory needed to store the interpreter, run-time system, and graphics. Estimates suggests that it would take approximately 14 megabytes of memory to hold only the course code on a MacIntosh microcomputer using Course of Action software. The future of the course in its present mode is not certain for a number of non-technical reasons, including maintenance costs on old hardware, curriculum changes, and the capabilities of microcomputers. Two tables depict the time requirements for students and computer memory requirements for each of the 14 topics covered by the one-semester introductory course. (4 references) (Author/EW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

research and information report

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

RIR-88-6

Fifteen Years of Teaching Elementary Applied Statistics Using CAI

S. Hunka
Faculty of Education
University of Alberta

division of educational research
faculty of education
the university of alberta
edmonton, canada

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Hunka

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

RIR-88-6

Fifteen Years of Teaching Elementary Applied Statistics Using CAI

S. Hunka
Faculty of Education
University of Alberta

November 7th, 1988
Division of Educational Research Services
Faculty of Education, University of Alberta
Edmonton, Alberta, Canada, T6J 2G5

Abstract

A CAI course in elementary applied statistics has been in operation for fifteen years in the Faculty of Education, University of Alberta. This course was to some extent an outgrowth of the development of a data processing facility which assisted staff and graduate students in the statistical processing of research data. The original development of the course was on an IBM 1500 system, and then moved to a DEC VAX 11/780 system in 1980.

Because numerous requests have been made about the nature of this course, frequently under the assumption that it can be easily moved to a micro computer, this document describes the nature of the course content as well as some of its technical characteristics.

The statistics course was originally created to be the primary mode of instruction, and therefore is very extensive in terms of content and style of presentation. The course includes fourteen "modules" of instruction, and ten examinations. It takes the average student about 80 - 90 hours of one-line instruction to complete the course. The range of times encountered for completion have been from 32 hours to 180 hours. It is estimated that at least 1500 graduate students have taken this course in CAI mode, and some general reactions of these students are noted.

The course was originally programmed in IBM's Coursewriter II authoring language for use on an IBM 1500 system. It continues to exist in this language, with some enhancements provided through the development of an in-house interpreter. Under the present CAI system, the course requires about 2.3 megabytes of memory not counting the memory needed to store the interpreter, run-time system, and graphics. Estimates suggest that it would take approximately 14 megabytes of memory to hold only the course code on a Macintosh micro computer using Course of Action software. An elementary explanation of why the course is not easily moved to other systems is noted.

The future of this course in its present mode of delivery becomes less certain for a number of non-technical reasons. These reasons include escalating maintenance costs on old hardware, curriculum changes, and capabilities of micro computers.

Fifteen Years of Teaching Elementary Applied Statistics Using CAI

S. Hunka
Faculty of Education
University of Alberta

Introduction

During the last five years with the widespread use of micro computers many new users of computer-assisted instruction (CAI) have been rediscovering its efficacy, the problems of developing courseware and making it operational, as well as learning about authoring languages, authoring systems, hardware, and software requirements. Few seem to realize that many current issues and problems were encountered at least fifteen years ago. Fewer still seem to realize, that even though hardware used for CAI fifteen years ago was technically much less advanced than today, student terminals with CRT screens, light pens, graphics, audio, and film projectors did exist.

Extensive courseware was operational in the late sixties and early 1970 primarily on the Plato system at the University of Illinois, and at as many as 25 other institutions and training establishments which made use of the IBM 1500 CAI system.

In the period between 1970 and 1980, the University of Alberta developed many lengthy courses (30 to over 100 hours average student time for completion) using the IBM 1500 system. Some of these courses migrated to a Plato or a DEC VAX 11/780 system after the University of Alberta returned its IBM 1500 system in 1980. Two courses, one in cardiology and the other in elementary statistics continue to be used in 1988, at least 15 years after their initial development. The cardiology course migrated to the Plato system, and more recently to the IBM PC using Infowindow, while the statistics course continues to be operational on a DEC VAX 11/780 system. A rough estimate indicates that up to 1988, each of these courses has been taken for course credit by at least 1500 different students.

Numerous requests have been made for information concerning the operation of the CAI system at Division of Educational Research Services (DERS), and about the statistics course. This paper serves to document from one of the course author's perspective the nature and operation of this courseware.

In The Beginning

In 1961 when the author joined the staff of the Faculty of Education, based upon his experiences in numerical computation at the University of Illinois using Illiac I, the process was started to develop statistical programs for use by educational researchers. A library of programs was developed which was used on the University's IBM 1620, later converted and expanded for operation on an IBM 7040 and IBM 360 series machines. The extensive statistical library was created much before such statistical packages as SPSS were available. As the importance of computers grew, it was recognized that although statistical programs assisted many staff and graduate students to complete their doctorates, there was little to support researchers who did not depend primarily upon statistics for their research. In order to support curriculum based research, an IBM 1500 system was ordered.

The IBM 1500 system was ordered at the same time an IBM 360 was ordered for general campus use. Key supporters in presenting the request to the Board of Governors were Dr. Max Wyman, the academic vice president, Dr. Don Scott, a physicist who was chairing a newly structured Department of Computing Science, and the author. The Board of Governors approved the acquisition of both the IBM 360 and IBM 1500. The delivery of the IBM 1500 was made in 1968 and placed on the ninth floor of the Education Building, south wing. Both computers were budgeted in the Department of Computing Science. A year later the operating budget for the IBM 1500 system was transferred to the Faculty of Education.

The IBM 1500 delivered to the University of Alberta was the first such system sited outside of the United States. Three or four years later, the Department of Quebec also obtained an IBM 1500 system, but it was operated for only about one year. DERS provided assistance to the Quebec installation by having one of its system analysts (Mr. N. McGinnis) spend a week with the Quebec-based group. During those years, it was politically more acceptable for Quebecois to travel to France than to Alberta, as numerous suggestions that Quebec send their operator for training to our installation were not accepted.

In support of the CAI installation in the Faculty of Education, the author received a $52,000 grant from the Donner Foundation (Canada). This grant was used to establish a graduate program in the area of CAI in the Department of Educational Psychology.

Prior to the delivery of the IBM 1500 system, DERS had developed a small data processing facility staffed primarily by graduate students interested in the area of tests and measurement. This facility included keypunches, an IBM optical test scorer, a card duplicator, a card sorter, and one of two remote terminal on campus to the IBM 360 which was available for limited use of APL. (The first terminal was located in the Department of Computing Science under the guidance of Professor Bill Adams, who had returned from working at Science Research Associates (SRA) in Chicago (SRA had then recently been acquired by IBM) and Dr. K. Smillie.) The University of Alberta's first delivery system of computer output from the computer center was established by DERS on a 6 days a week basis. During this period the Faculty of Education used about 10% of the computing center's resources.

In about 1967, the first patient management simulation (a hypertension patient) was created at DERS using APL. It was the latter experience, and the experience in numerical programming which established the concept of separating curriculum content from courseware programming code, later to be conceptualized by Dr. E.W. Romaniuk in his doctoral thesis. At the time this work was being done, the author had a joint appointment (Honorary) in the Department of Internal Medicine, then chaired by Dr. Don Wilson. At that time the Faculty of Medicine was very actively involved in curriculum revision, and Dr. Wilson was actively involved in the examination system established by the Royal College of Physicians and Surgeons. Under the forward looking vision and the impetus provided by Dr. Wilson, the R.S. McLaughlin Examination and Research Centre was established at the University of Alberta for the creation and scoring of Royal College speciality examinations. The R.S. McLaughlin Centre initially drew upon the data processing facilities of DERS, and later upon the IBM 1500 system for the further development of patient management examinations.

The development of courseware on the IBM 1500 system required detailed programming in IBM's Coursewriter II (IBM, 1970) authoring language. One of the first CAI courses was created by Dr. Brian Hudson who was on sabbatical leave in the Faculty of Medicine from the University of Monash in Australia. Dr. Hudson acquired sufficient skills,

among many of his other teaching and research duties at Alberta, to be able to "cold start" and operate the IBM 1500 system by himself. Dr. Hudson was the only academic staff member, outside of DERS, who was ever given this privilege. The work by the Faculty of Medicine was actively supported by Dr. Walter MacKenzie, Dean of Medicine who frequently visited the Division, and later was to provide support against budget cuts assessed against the Faculty. It was this initial work in the Faculty of Medicine which later established what was likely the world's first computer-based medical examination which was run on the same day on a national basis. During this time considerable support and encouragement was provided by Dr. M. Horowitz, Dean of Education.

In an attempt to demonstrate to the Faculty of Education the potential of the IBM 1500 system, two courses were started. Dr. T. Maguire started a course which introduced the student to analysis of variance, and the author created a course to teach Coursewriter II programming. To further demonstrate to the Faculty how the IBM 1500 system could be used, a course in elementary statistics was started, and the initial work of Dr. Maguire was incorporated into the present course.

It has been estimated that 2500 man hours were required for the initial development of the statistics course (Hunka, Romaniuk, Maguire, 1976). The initial development costs were estimated to be $10,000, and did not include the costs associated with the contributions made by the academic staff members. It is now estimated that the total development time, including the conversion of the course to operate on the VAX system to be about 3500 man hours.

The Statistics Course: Content

The content of this course generally fits the content of a first course in statistics for graduate students in educational psychology who have had little or no previous courses in statistics. The course is most frequently taken by graduate students in educational psychology, physical education, and nursing. The statistics course currently contains the following major modules:

Descriptive Statistics: The following topics are included: grouping data to form a frequency table, class limits, the histogram, area contained by a histogram, area as it relates to probability, measures of central tendency, the arithmetic mean minimizing sums of squares, variance and standard deviation, deviation and Z-scores, introduction to the normal curve; effects of linear transformations on the mean and variance. An introduction to the stem and leaf, and box and whiskers representations of data is made. There is a 56 item test associated with this section. (Author: S. Hunka)

Normal Curve and Binomial: This section is devoted to calculation of areas under various portions of the normal curve based upon tables contained in most statistical text books. The binomial is introduced through an examination of the number of different ways that 5 test items scored right/wrong can produce total scores of 0 to 5; combinations and factorial is introduced. The equation for calculating the probabilities is used and Pascal's triangle is introduced. There is no examination associated with this section. (Author: S. Hunka)

Prediction: Prediction is introduced through the use of a scatter diagram and a regression line fitted to minimize the sum of squares of the errors of prediction. A calculus approach (optional) is taken to define the constant "a" and the regression coefficient "b" in the equation y=a+bx+e. Both regression lines are introduced, i.e., the line to predict y from x, and x from y. The effect of linear transformations on "a" and "b" are examined. The usefulness of plotting residuals is introduced. There is a 33 item test for this section. (Author: S. Hunka)

Correlation: The Pearson product moment correlation is introduced, initially using the regression coefficients covered in the prediction section. Stress is upon the correlation defined as the ratio of the covariance to the product of the two standard deviations. The proportion of "variance accounted for" is introduced. The effect of linear transformations upon the correlation coefficient is also covered, as well as conditions under which the coefficient cannot be defined, e.g., when one variable has a zero variance. Scatter diagrams are related to the correlation. A 31 item test is also available for this section. (Author: S. Hunka)

Other Types of Correlation Coefficients: This section covers other types of correlation coefficients, specifically the biserial and point biserial correlation, and the Phi coefficient. Those which can be derived from the Pearson product moment correlation are noted. A 23 item test is available for this section. (Author: S. Hunka)

Partial & Multiple Correlation: The usual partial correlation coefficient is introduced from a simple algebraic point of view. The multiple correlation coefficient is introduced algebraically, as well as through the use of the inverse of the correlation matrix. (The student is not introduced to matrix algebra, and is only presented with how to use the inverse matrix). The effects upon the calculation of the squared multiple correlation of a non-gramian correlation matrix are considered. (Author: S. Hunka)

The Random Sampling Distribution of the Mean: The student is introduced to the terms "statistic", "population parameter" and sampling, the sampling distribution of the mean, the expected value of the mean, and the variance of the distribution of the means. (Author: T. Maguire)

Hypothesis Testing: The student is introduced to hypothesis testing including the alpha and beta errors, as well as power. This is done in the context of the one sample case involving a sample mean. The student is expected to be able to calculate beta, and power and to understand the effects of sample variability and sample size on these calculations. A 22 item test is available for this section. (Author: E.W. Romaniuk)

Statistical Tests: The following statistical tests and related confidence intervals are covered: t-test for dependent and independent samples, using the t table, confidence intervals; t tests involving correlation coefficients, Fisher Zr transformation, and confidence intervals; one-tail and 2-tail hypothesis testing; tests involving proportions. A 23 item test is available. (Author: S. Hunka)

One-Way Analysis of Variance: This section is divided into four subunits: Unit 1 covers the purpose of the 1-way anova and introduces related terminology. Unit 2 covers computational aspects which include calculations for various sums of squares and mean squares. Unit 3 presents an intuitive rationale for the anova based on the expectation that the mean square between will reflect mean differences among groups while the mean square within will not. Unit 4 which is optional covers the mathematical rationale for a 1-way anova. An 18 item test is also available for this section. (Author: T. Maguire)

Tukey and Scheffe Tests: Tukey and Scheffe approaches for testing contrasts of means is introduced. The use of the studentized range statistic table is introduced for the Tukey test, while the use of orthogonal contrasts is introduced for the Scheffe procedure. A 15 item examination is available. (Author: T. Maguire)

2-Way Analysis of Variance and Interaction: The stress in this section is the interpretation of interaction based upon plots of mean performance vs treatment. Although the 2-way anova is also introduced the computational aspects of various sum of squares is not stressed. A 20 item test is available for this section. (Author: T. Maguire)

Chi-Square: The use of Chi-square for testing goodness of fit and its use with contingency tables is covered. A 23 item test is available for this section. (Author: E.W. Romaniuk).

Analysis of Research Data Using a Computer: This section covers such topics as the steps frequently used in carrying out a research project, nominal, ordinal, interval and ratio scales of measurement, selecting statistical parameters to reflect treatment effects, a cursory introduction to computing; the construction of Fortran format statements. A small portion of this section is unique to the preparation of records used with the University of Alberta's Amdahl computing system using MTS. (Author: S. Hunka)

The topics are listed in the order in which a student would normally study them, however the current CAI run-time system includes a registration file for each student

within which the sequence of the topics can be specified. If a section contains an examination, the examination immediately follows the instructional section.

Based on experience with graduate students who have taken the course prior to entering a graduate program in the United States, upon providing the contents of the course to the U.S. institution, successful completion of the course has always been accepted for credit.

The Time Requirements of the Course

The extensiveness of the course should suggest that it is not a trivial course so frequently encountered in computer-based instruction, e.g., a course of length measured in minutes or 1 to 2 hours. Table 1 lists the times in hours which the "average" student might expect to spend in covering each section of the course.

Table 1 Expected Time to Cover Topics of the Course

Section	Content	Time(Hrs)
0	Introduction to Terminal	.5
1	Descriptive statistics	6.0
2	Normal Curve & Binomial	5.0
3	Prediction	6.0
4	Correlation	5.0
5	Other Types of Correlation	4.0
6	Multiple & Partial correlation	4.0
7	Inferential Statistics	2.0
8	Hypothesis Testing	6.0
9	Statistical Tests	8.0
10	1-way Analysis of Variance	8.0
11	Tukey & Scheffe Contrasts	7.0
12	Interaction	5.0
13	Chi-Square	2.0
14	Data Analysis	4.0
	Total Time in Hours	72.5

The times noted in Table 1 do not include time for completion of the on-line examinations nor the time a student might use in repeating a section. The course is currently a one semester course.

Student Use of the Course

The course was originally written with the objective that it should form the primary source of instruction, and that the instructor should intervene when the computer-based instruction is not adequate for a student. In such a mode of presenting this course, one student may receive 3 hours of individual help from an instructor, while another student might receive 15 hours of individual help. These differences would reflect differences in the backgrounds of students and the ease with which they handle arithmetic and algebraic concepts. In addition to the computer-based instruction, students are expected to complete ten laboratory assignments. These have usually

involved calculations which paralleled the major sections of the course, and were expected to be done using a small pocket-type calculator. Some instructors have introduced the use of APL or SPSS as a part of the laboratory assignments.

Students who have tended not to do well in the course can be characterized as follows:

1) The older student (45 years and above) who may be quite facile at arithmetic but has a poor understanding of arithmetical concepts.

2) The student who cannot handle elementary operations using fractions, e.g., obtaining a common denominator before adding fractions.

3) Not understanding the difference between 0+1 and 1+0.

4) Unable to expand the term $(a+b)^2$.

5) The student who is highly anxious when handling arithmetic, and cannot make elementary estimations, e.g., cannot distinguish the relative sizes of 1+10 and 1+.01 These students tend to be overly dependent upon results from a pocket calculator, but cannot determine whether the calculator is giving a reasonable result.

6) A student who does not set a consistent pace in covering the course under the assumption that "cramming" at the end of the semester will be sufficient, tends to encounter problems. It is not possible to "force" oneself through the course. Wrong answers will simply prevent the student from moving quickly through the course.

There is generally a slightly negative correlation, e.g., -.2 between "time on system" and the final grade in the course. Students who do not get 46 or more on the first test can be divided into two groups, one group which has poor understanding of arithmetical concepts and a second group which, although having reasonable conceptual understanding of arithmetic, have not made use of these for a long period of time. The former group cannot easily recover from their problems. Unfortunately, many of these students allow their learning at the terminal to deteriorate to copying everything off the terminal screen into their notes. This activity may produce very detailed notes, but the activity is simply that of copying with little understanding.

When the computer-based examinations were counted towards the final course mark, a small percentage of the students, e.g., 3%, would appear to spend an inordinate amount of time agonizing over the answer to a question. After two or three incorrectly answered questions, tears would result. A small number of students have complained that a question was marked "incorrectly", e.g., they typed the correct choice, but the computer did not recognize the choice. The latter problem rapidly disappeared when the student was shown that a complete record of their input could be recovered from the records saved by the system.

During the last few years instructors have encouraged students to take the on-line examinations even though the results have not been counted towards a final course mark because security of the examination questions has been lost. After fifteen years every question and its answer has been copied by students.

Students for whom the CAI presentation is advantageous are those who can progress quickly through the course and complete the course earlier than the scheduled time, students who have a language problem can compensate by using a slower pace commensurate with their understanding of English, and those students who wish to proceed at a slower pace with quite frequent review of the course content. The course, when used with summer session students, has allowed the better students to complete the course in four weeks rather than in six weeks. As well as providing the better student with this advantage, students who have anticipated difficulty with the course have been allowed by some instructors to start the course before the formal starting date of the summer session. During summer session, an 8 to 10% failure rate has been typical.

The course has been used by a small number of deaf students who successfully completed the course with respectable marks. Parenthetically, when working with deaf students it is desirable for the instructor or teaching assistant to be able to quickly use the terminal as a basis of communication with the deaf student.

More recently, students who cannot attend regular lectures during the day because they have a full time job, or baby sitting constraints, find it advantageous to use the CAI system at times convenient to their special schedules. The latter has become more important as the number of female students taking the course now substantially out numbers the male students. Students should, however, have access to an instructor or teaching assistant at regular intervals in order to have any minor misconceptions clarified and to verify that their progress is satisfactory. In the 1970's the CAI facility operated 14 hours per day, Monday to Friday, 8 hours on Saturday, and 4 hours on Sunday, providing considerable opportunity for students to use the system. Currently the CAI facility is only operational 8 hours per day on week days only.

In reviewing our experience with students during the last fifteen years, one suspects that for a majority of students their numerical skills are poorer today than fifteen years ago. Two sections of the course have optional routes available if the student has an elementary understanding of calculus. Today, this is not a very meaningful option for most students. Ten years ago, almost all students taking this course had a course in tests and measurement in their undergraduate education program. This assured that the students would have had some contact with such concepts as the mean, standard deviation, correlation, standardized scores, etc. During the last ten years, the Faculty of Education no longer requires undergraduate students to take such a course, thus, students progressing to graduate work are not likely to have encountered these concepts. This phenomena seems to be part of a more general trend away from formal and analytic approaches in our Faculty of Education.

A cursory examination of the course may suggest, depending upon what section is being viewed, that there is not much student interaction or responding. Experience suggests that a typical student will leave 1500 to 2000 performance records, i.e., records detailing the nature of student input to questions posed. This number does however, include repetitions such as would be made when a student repeated a section or an examination. This is considerably more interaction than would be observed using the traditional lecture.

It would be incorrect to suggest that every student who has encountered the course has enjoyed the CAI mode of delivery. The reaction of students is in part related to the manner by which the instructor participates in the delivery of the course. Some instructors have attempted to make use of the course without actually having gone through the course themselves! Under such circumstances, the instructor cannot quickly diagnose a conceptual problem for the student at the terminal. Nevertheless, even with

such instructors, it is assured that the students will cover the curriculum of the course. This can be an advantage to an institution which cannot guarantee the availability of qualified instructors. Instructors who use the CAI courseware as a basis for presenting the major components of a course, and then add additional components on topics of special interest to themselves can also create problems. The course simply begins to double in size and students complain about the workload.

Complaints voiced by students, usually the poorer students, include inability of the CAI system to provide a frame-by-frame backing-up facility, unavailability of course notes preferably created from each screen display, left or right cerebral dominance hampering reading from the screen, learning dysfunctions, insufficient availability of access to the system, e.g., evenings, and so on. Unfortunately, seldom have students who have done well in the course taken the time to voice their approvals. Although a set of notes based on the screen displays could be made, the task would be very large indeed producing thousands of pages. Of course, as is true in other universities, statistics is an anxiety producing course for most graduate students in education. It is likely for most students who have taken this course in computer-based mode, that this mode was the lesser of two evils.

Technical Characteristics of the Course

With the widespread use of micro computers, requests have been frequently received for a copy of the courseware for purposes of placing it on a micro. Most such requests are made by instructors who are quite naive concerning the portability of courseware and related memory requirements, run-time systems, and operating systems. Table 2 gives the amount of memory in bytes (characters) for each major section of the course, but excluding the graphics requirements.

Table 2 shows that a total of 2.3 megabytes of memory is required to store the course. On our current CAI system, this number is the memory requirement to store the source program code, which is directly interpreted by a Coursewriter II interpreter. This number does not include the graphics code nor the code for the interpreter and run-time system. There are approximately 350 diagrams (graphics) which must be stored in digital form.

Table 2 Memory Requirements for the Statistics Course

Section	Content	Bytes
0	Introduction To Terminal	27136
1	Descriptive statistics	195072
	examination	45568
2	Normal Curve & Binomial	199168
3	Prediction	145920
	examination	28160
4	Correlation	56320
	examination	27648
5	Other Types of Correlation	113664
	examination	22016
6	Multiple and Partial Correl.	103936
7	Inferential Statistics	81920
8	Hypothesis Testing	132608
	examination	32256
9	Statistical Tests	134656
	examination	25088
10	1-Way Anova	149504
	examination	22016
11	Tukey & Scheffe Tests	159232
	examination	19456
12	Interaction	236032
	examination	18944
13	Chi-square	39424
	examination	20480
14	Data Analysis	217088
	Total memory in bytes:	2253312

The graphics are relocatable under our run-time system. Systems which cannot make repeated use of a graphic, would require considerably more storage requirements. Recent attempts to convert the first section of the course (descriptive statistics) by using the Course of Action software for running on a MacIntosh micro computer, suggests that the complete course would require approximately 14 megabytes of storage! Although such storage is available with hard drives, devoting such a large amount of memory to one course effectively reduces the use of the micro to one course. Of course "future" use of removable compact disks would get around this problem, however, the course would become permanent in such a storage medium.

The Instructional Strategies Used in the Course

If the instructional strategy had to be characterized by a single strategy, the course would be described as of a tutorial nature. This however, would not truly reflect the nature of the course. There is no single instructional model which has been used throughout the course. The nature of any instructional paradigm was created by an author on the basis of the instructional task at hand. The following instructional strategies are used in various sections of the course:

1) Tutorial, e.g., a section which requires the student to gain an understanding that an infinite number of normal curves exist is done on an inferential approach through recognizing that one curve could exist for each value of the mean and variance selected.

2) Diagnosis and remediation, e.g., diagnostic quizzes are given at various intervals. These diagnostic quizzes are different than the section examinations in that they have considerably more prompting, hints, and attempts to track the steps the student has used to obtain an answer. In the Interaction section, a student's error is identified with respect to an error in ascertaining row effects, column effects, and interaction effects. The remediation resulting from such errors is a function of the combination of errors which were made.

3) Drill and practice, e.g., in the section on using the Normal Curve table to determine areas under the normal curve, the student is asked to make many calculations of the area under different parts of the curve using raw score or Z-score values.

4) Calculations, e.g., the student is asked to make numerous calculations such as for example completing a table of probabilities for a binomial distribution, calculating a correlation coefficient, Chi-square value, etc.

5) Simulation, e.g., a student is presented with a series of unordered steps for the calculation of a squared multiple correlation. The student must select each step in its appropriate order, and at each stage is presented with information which must be assessed as being correct or in error. In studying the binomial, the student is asked to guess the order of a sequence of nonsense words which the computer is "guessing". The probability of the student guessing the same sequence as guessed by the computer is calculated.

6) Testing, e.g., ten tests are linked to follow most of the instructional sections. An answer is immediately scored for the student and the correct answer displayed. The response to each question is stored and can later be retrieved for discussions with the instructor. The final examination scores as they are accumulated by the system are always available to the student.

Different strategies were required because instructional tasks vary considerably throughout the course. In addition, experience suggests it can be extremely boring for a student to be presented with the same instructional strategy for eighty or more hours of instruction.

Implementation: Technical Aspects

The courseware was implemented on a Digital Equipment VAX 11/780 computer with Gigi micro-terminals and a Barco color monitor. The IBM 1500 implementation of the course used a terminal with capabilities for displaying varying character fonts designed by the authors, a black and white screen with graphics capability, a light pen, and an image projector. At minimum, all these capabilities had to be available for the VAX implementation. In addition, since terminals with color capability were available with the VAX, enhancements which come with the use of color were incorporated.

The interpreter for the Coursewriter II authoring language in which the course was written, and the run-time system was written using a special language called Elf

(Education Language Facility) developed by Davis (Davis, 1985) at the Division of Educational Research Services. This language was specifically designed to allow the easy creation of CAI application software such as interpreters and compilers. It provides for easy use of complex data structures and has excellent parsing capabilities. Some components of the Elf language are machine dependent, however most software components are machine independent. Thus, Elf was designed for easy movement to other computing systems. The Elf system, the Coursewriter interpreter, and the run-time system, all operate under the VAX VMS operating system.

In moving the statistics course from the IBM 1500 system to the VAX system only slight modifications (usually enhancements) were made to the Coursewriter source code. Such modifications were possible because the interpreter was under local control rather than under the control of a supplier. The following are examples of the enhancements which were made:

1) Although the "dt" (display text) of Coursewriter is still acceptable, to improve I/O to the screen a "db" (display block) was created.

2) In order to accommodate the equivalence of an image projector, three screen formats were defined rather than the one, 40 column screen format used with Coursewriter. Codes, sv0, sv1, and sv2 cause the interpreter to create a 40 column screen, a split screen with 40 columns on the left-half of the screen, and an 80 column screen respectively. The right half of the split screen format is used to display a drawing which would have normally been displayed by the image projector on the IBM 1500 system. Codes used in Coursewriter to search for and display an image, close and open the shutter were maintained, but their action were now in the context of a digital image rather than a photographic image.

3) The definition of color was added to the parameters associated with the display of text.

4) The interpreter would accept Regis code as well as Coursewriter code. This is used, for example to create such features as a box around a title.

5) The Coursewriter code which activated the light pen, now causes the Gigi cross-hairs to appear on the screen.

6) Special codes for cursor control, e.g., backspace, reverse index, index, font change, were allowed within a text stream.

7) A special code was created which allowed the author to enable or disable branching by the student.

8) Label statements could now contain a comment identifying the nature of the curriculum content to follow.

9) The response identifier associated with the request for a student input could now be used to cause the automatic storage of a final examination mark.

10) Coursewriter II program codes to display a film image are still used, but reference is now made to a file containing DEC's Regis code for each specific image required.

The CAI Run-Time System

Frequently current users of micro-based CAI do not understand that courseware must have a run-time system under which it operates. Functions of the run-time system include the following:

1) a performance recording system to record keyboard and pointing responses. Responses must be associated with specific locations in the course, and date stamped in order that they can be meaningfully interpreted.

2) a method is required for restarting the course for the student at a point in the course which is logical from a curriculum point of view, and which ensures the screen is properly created and formatted.

3) allow access to the proper course files including the library of graphics.

4) link the course segments together according to a "routing vector" specified in the student's registration file.

5) control legitimate student movement to different parts of the course or within a course section using a table of contents.

The statistics course operates under a run-time system which provides, in addition to the items mentioned above, a host of support and control functions during the execution of the course.

In addition to the run-time system supporting the student and the courseware requirements, it must also support some functions for the instructor. An instructor is usually interested in monitoring student progress, whether it be examination results, or identification of where the student is in the course and how much time the student has spent taking the course on the system. The run-time system under which the statistics course operates, allows, for example, the instructor to retrieve examination records detailing the response to each question. Even though the examination questions are randomly presented to the student, the retrieval allows the instructor to have the question, the student's answer, and the correct answer to be displayed. The instructor can then discuss the question with the student, enter an answer and have the system score the answer entered. The instructor can also enter non-computer-based examination results, which can then be displayed with the computer-based examination results by the student at will. Thus, it has been possible for an instructor to enter marks from home via a remote terminal, and have the marks immediately available for the students the next day.

Running the Course on Another VAX System

The statistics course and the associated support systems have been operated on VAX machines off the University of Alberta campus. The statistics course was operational in Fort McMurray, Alberta, during the winter session of 1988. The course has also been made operational for viewing purposes at Dalhousie University, and is currently operating as an adjunct to lectures in the Faculty of Education at the University of Calgary. The course and system has also been provided to the University of Saskatchewan, but it is unknown whether the system has been made operational. The CAI system at DERS and the statistics course has been successfully run experimentally on a MicroVax.

Experience suggests that some small VAX installations do not have technical personnel who fully understand the parameters of the machines which they are operating. Thus, problems are encountered in bringing the course and system up. In addition, care must be taken to ensure all student files are properly defined, and logical file assignments correctly made. These aspects of making the course operational cannot be done haphazardly.

Transferring the Statistics Course To Another Computing System

Moving the statistics course to another system cannot be accomplished by an instructor of statistics simply by having a copy of the source Coursewriter II code. The following are a few of the constraints which would have to be considered in moving the course to another machine:

1) An Elf interpreter for the Coursewriter code would be required. Although Elf was designed to be portable to another machine, some specialized assembler code would be required to interface, for example, input/output requirements to the host machine. Although our CAI system and the statistics course has been run on a MicroVax, it has not yet been ported to a totally different machine.

2) The output to the screen display from the Gigi terminal assumes a screen with 768 by 480 pixels in the colors blue, red, magenta, green, cyan, yellow, and white.

3) The output to the student's terminal is in DEC's Regis graphic code. Experience has suggested that many software emulators, which are claimed to interpret Regis code for a terminal other than a DEC Gigi terminal, simply do not handle all the codes correctly. Problems occur most frequently in displaying the diagrams used in the statistics course.

Course Evaluations

No recent formal evaluation has been made of the course. One evaluation was conducted in 1975-76. No comparisons have ever been made between student achievement based on this course and regular classroom lectures. Average examination grades in 1976 on ten topics of the course ranged from 73% to 93%, with the lower mark being for the section on data analysis and research design, while the higher mean being for Chi-Square (N=32). The majority of examination marks were found to be between 75 and 90 percent. Without being able to relate the examination questions to some more standardized measure, it is of course difficult to judge their validity. The most difficult sections of the course as reported by students in 1976 were the sections covering statistical tests and the intuitive and mathematical rational for the analysis of variance. The easiest sections were reported to be terminology used in inferential statistics, the purpose of the 1-way analysis of variance, and Chi-Square. Students rating the course on a 7 point Likert-type scale suggested that the topics of the course were easier to learn by CAI than in the regular classroom. The mean rating for the statement measuring this characteristic of the course was 5.6.

In 1979 another evaluation was conducted by Vargo (Vargo,1979). This evaluation identified 117 students who had taken the course between 1976 and 1978. These

students were requested to complete an attitude questionniare and a 15 item achievement test. Unfortunately, the return rate for the questionnaire was only 47% and therefore there is doubt as to the validity of this assessment. Vargo concludes that

> ".... STAT1 has proven to be regarded positively by most students with the drawbacks being far outweighed by the advantages. Computer-assisted instruction is a viable alternative to classroom instruction and, while the results of this study are not intended to indicate that students prefer CAI over classroom instruction in all instances, they certainly suggest that students prefer to learn statistics via CAI."

The Future of the Statistics Course

Although extensive effort has been placed into the development of this course and its movement to the VAX system, there is no assurance that the operation of the course will continue in CAI mode. There are many factors which relate to this situation. These are discussed below.

The development of applied statistics has not stood still. Although some new material has been added to the course, e.g., a very brief introduction to such data representations as stem and leaf, box and whisker diagrams, and plotting of residuals, the basic content has remained relatively static. Some instructors feel that more emphasis should be given to exploratory data analysis.

Some instructors desire to base their applied statistic courses on the running of computer programs which form a basis for class discussions of the output, while some others feel that the use of spread sheets enhances the students understanding of applied statistics. The statistics course described here uses neither of these forms of instruction. It is the author's opinion that running computer programs followed by discussions of the results would provide meaningful laboratory assignments. To some extent so would the use of spread sheets. We question whether the use of spread sheets would allow the student to generalize computational procedures to symbolic form. For a more comprehensive understanding of computational procedures, it would seem that APL would be a more powerful tool and lend itself to better generalizations than the use of the obtuse spread sheet statements required to carry out arithmetic operations. These alternate approaches to teaching suggest different instructional objectives.

Some researchers in the social sciences feel that no meaningful measurements can be made on human subjects in the educational context. For some, the only useful research approach is ethnography. More generally, the trend in education away from formal analytic approaches works against teaching elementary applied statistics. An intermediate position is voiced by the statement that researchers do not have to understand the computational aspects of statistics, they need only to have access to a program which provides the statistical results.

Even though the statistics course described here was written with a view to it being the primary source of instruction, an instructor or teaching assistant is required to back the system up for some students. Today, the cheapest manner of course delivery is to hire an "academic grape picker", i.e., a sessional instructor, and fill a classroom with fifty students. (Parenthetically, in using the term "academic grape picker" we do not wish the term to be derogatory towards sessional instructors. Unfortunately, these

people have been caught in the politics of university funding policies of governments.) The maintenance costs of the VAX 11/780 increase each year, and are reaching a size which raises concerns, particularly relative to human resource needs. Here we would acknowledge, unlike IBM during the operation of the 1500 system, DEC has taken no interest in these problems, nor our software and courseware developments.

There is absolutely no question that there is an important role for micro computers to play in computer-based instruction. We anticipate that micros will grow more powerful as will main-frame computers. Eventually, the statistics course which we have operational on a VAX 11/780 could be made operational on a micro computer. Perhaps making it operational is a function of acquiring the human resources needed for such a project, perhaps it is a function of waiting for CAI software on micros which would accommodate the present features of the course and its run-time environment, perhaps it is a matter of throwing out features of the courseware and run-time system and declaring them unnecessary.

References

Davis, A., Elf Reference Manual, Version 5 Level D, Division of Educational Research Services, University of Alberta, Edmonton, 1985.

International Business Machines, IBM 1500 Coursewriter II Author's Guide, Programming Publications, San Jose, Calif., 1970.

Hunka, S., Romaniuk, E.W., and Maguire, T., Report on the Use of the Computer-Assisted Course STAT1 as used in Educational Psychology 502, RIR-76-7, Division of Educational Research Services, University of Alberta, 1976.

Vargo, F.A. Follow-up Evaluation of STAT1 Computer Assisted Instruction Program at the University of Alberta, RIR-79-4, Division of Educational Research Services, University of Alberta, Sept. 1979.